Satellite Precipitation Characterization, Error Modeling, and Error Correction Using Censored Shifted Gamma Distributions

Daniel B. Wright*

Civil and Environmental Engineering

University of Wisconsin

Madison, WI

Dalia B. Kirschbaum

Goddard Space Flight Center

National Aeronautics and Space Administration

Greenbelt, MD

Soni Yatheendradas

Goddard Space Flight Center

National Aeronautics and Space Administration

Greenbelt, MD

Earth System Science Interdisciplinary Center

University of Maryland

College Park, MD

*danielb.wright@wisc.edu

Abstract

Satellite multisensor precipitation products (SMPPs) have a variety of potential uses, but suffer from relatively poor accuracy due to systematic biases and random errors in precipitation occurrence and magnitude. We use the Censored Shifted Gamma Distribution (CSGD) to characterize the Tropical Rainfall Measurement Mission Multi-Satellite Precipitation Analysis (TMPA), a commonly-used SMPP, and to compare it against the rain gage-based North American Land Data Assimilation System Phase 2 (NLDAS-2) reference precipitation dataset across the conterminous United States. The CSGD describes both the occurrence and the magnitude of precipitation. Climatological CSGD characterization reveals significant regional differences between TMPA and NLDAS-2 in terms of magnitude and probability of occurrence. We also use a flexible CSGD-based error modeling framework to quantify errors in TMPA relative to NLDAS-2. The framework can model conditional bias as either a linear or nonlinear function of satellite precipitation rate and can produce a "conditional CSGD" of describing the distribution of "true" precipitation based on a satellite observation. The framework is also used to "merge" TMPA with atmospheric variables from Modern-Era Retrospective analysis for Research and Applications (MERRA-2) to reduce SMPP errors. Despite the coarse resolution of MERRA-2, this merging offers robust reductions in random error due to the better performance of numerical models in resolving stratiform precipitation. Improvements in the near-realtime version of TMPA are relatively greater than for the higher-latency research version.

1. Introduction

Precipitation data is critical in a variety of subjects including climate studies, meteorology, hydrology, and natural hazards. While precipitation is relatively easy to measure at a single point using a rain gage, measurement over large regions at high spatial and temporal resolution is a major challenge. A "constellation" of earth-observing satellite missions, including the now-defunct Tropical Rainfall Measuring Mission (TRMM) and the follow-on Global Precipitation Measurement (GPM) mission, co-led by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency. These satellites provide a mix of direct measurements of precipitation and related processes using active radar and indirect measurements using passive microwave (PMW), and infrared (IR). Satellite multisensor precipitation products (SMPPs) merge these various observations to create near-global precipitation records that approach two decades in length. Examples include the 3-hourly, 0.25° Tropical Rainfall Measurement Mission Multi-Satellite Precipitation Analysis (TRMM TMPA; Huffman et al., 2010, 2007); the 30-minute, 8 km Climate Prediction Center (CPC) Morphing Technique (CMORPH; Joyce et al., 2004); and the hourly, 4 km Precipitation Estimation from Remote Sensing Information using Artificial Neural Networks (PERSIANN; Sorooshian et al., 2000). Most SMPPs are available in near-realtime (with latency on the order of several hours) and some have non-realtime variants that utilize ground-based rain gage information for bias correction. Launched in 2014, the GPM mission builds on TRMM's legacy with an advanced active and passive instrument package. NASA's 30-minute, 0.1° Integrated Multi-satellitE Retrievals for GPM (IMERG; Huffman et al., 2014) dataset builds on more

than a decade of experience with SMPPs, combining the strengths of TMPA, CMORPH, and PERSIANN and incorporating additional improvements.

Despite widespread interest in SMPPs, these datasets often exhibit considerable errors, both systematic (i.e. bias) and random, stemming from a variety of sources. Observation quality varies within the satellite constellation, with active radar being the most accurate, followed by PMW and IR. Sensor technology and resolution varies with age and mission. The current constellation of satellites provides a PMW observation for most locations on Earth approximately every three hours, while radar observations are much less frequent. Between PMW measurements, algorithms typically use spatiotemporal interpolation of PMW or "infilling" using lower-accuracy IR. PMW observations tend to be more accurate nearer the tropics and for convective than for stratiform storm systems (Ebert et al., 2007) and are influenced by the underlying land or water surface, and microwave emissions from snow or ice-covered ground can be difficult to distinguish from emissions due to ice scatter in precipitating clouds (Ferraro et al., 2013; Ringerud et al., 2014; Tian and Peters-Lidard, 2007). IR and PMW instruments have difficulties with orographic precipitation systems due to their shallow nature (Shige et al., 2013) and high variability in microscale and macroscale dynamics (Anders et al., 2007).

Given the potential usefulness of SMPPs, it is natural to want to characterize SMPP errors using an error model that compares SMPP against "ground truth," i.e. more reliable reference data (typically rain gages or ground-based weather radar). Systematic error is usually heteroscedastic (i.e. depends on precipitation observation magnitude), a

phenomena known as conditional bias (Ciach et al., 2000). Such errors tend to be multiplicative (Tian et al., 2013) with a magnitude that increases with precipitation observation intensity. Error models can be used to identify and thus remove systematic errors. They can also describe the statistical distribution of random errors, which can be understood as the residuals once the systematic error has been removed. Using this approach, individual random errors are irreducible without some sort of additional explanatory information.

SMPP characterization efforts (e.g. AghaKouchak et al., 2011; Behrangi et al., 2011; Tian et al., 2009) often distinguish between three error "cases": false alarms, in which the SMPP reports precipitation while the reference data does not; misses, in which the reference reports precipitation while the SMPP does not; and hits, in which both report precipitation, but not necessarily of the same magnitude. Most error models that have been developed in the context of precipitation estimation using ground-based radar (AghaKouchak et al., 2010; Ciach et al., 2007; Germann et al., 2009) and SMPP (Gebremichael et al., 2011a; Sarachi et al., 2015; Yan and Gebremichael, 2009) have tended to focus on hit cases only.

Several previous SMPP error models have considered false alarms, misses, and hits separately, and then recombine these separate descriptions to create an overall estimated distribution of true precipitation. For example, the Precipitation Uncertainties for Satellite Hydrology framework (PUSH) introduced by Maggioni et al. (2014) uses a Gamma distribution to describe the precipitation intensity associated with misses, exponential decay and linear regression models respectively to describe the probability and intensity

114 associated with false alarms, and a generalized linear model to generate a Gamma
115 distribution of precipitation magnitude associated with hits. PUSH also uses a uniform
116 distribution to describe possible trace precipitation associated with cases where neither the
117 SMPP nor reference data report precipitation. For any zero or nonzero SMPP observation,
118 a probability distribution can be generated by combining these cases. The two-dimensional
119 Satellite Rainfall Error Model (SREM2D) introduced by (Hossain and Anagnostou, 2006)
120 takes a somewhat similar approach, but incorporates spatial and temporal autocorrelation
121 functions to construct ensembles of correlated precipitation fields.
122
123 This study applies a new shifted gamma distribution (CSGD) methodology to characterize
124 precipitation and create an SMPP error model that produces a "best guess" distribution of
125 the true precipitation by considering hits, misses, and false alarms. The CSGD technique
126 presented in this paper is arguably simpler than most, and comparison with the PUSH error
127 model that suggests that this relative simplicity is advantageous.
128
129 Previous precipitation error model studies have generally focused on relatively small
130 geographic areas where spatial stationarity of rainfall and model parameters can be
131 assumed; however, these approaches have not explored spatial variability in these
132 parameters or in model performance. This study is one of the few, along with Maggioni et
133 al. (2016), that applies an error model over a large region to better understand SMPP
134 performance characteristics and how they are tied to physiographic and climatological
135 features.
136

This study moves beyond the traditional notions of precipitation error modeling towards error correction by allowing the incorporation of additional information to reduce random errors. Previous researchers have suggested that topography and other land surface characteristics as well as other atmospheric variables such as humidity could help understand and, in principle, correct SMPPs (Gebregiorgis and Hossain, 2013; Gebremichael et al., 2011a). As far as we are aware, this study is the first to explore the potential benefit of incorporating atmospheric variables such as humidity and precipitation from numerical weather models (specifically, atmospheric reanalysis) in a satellite precipitation error model to reduce SMPP random errors. This is a promising approach since the complementary performances of numerically-simulated and remotely-sensed precipitation estimates provide the opportunity to produce merged datasets with smaller systematic and random errors.

The SMPP, ground reference, and atmospheric reanalysis datasets utilized in this study are described in Section 2. The CSGD and the CSGD-based precipitation error modeling and correction frameworks are introduced in Section 3. Results for precipitation characterization and SMPP error modeling are provided in Section 4. Summary and closing discussion follow in Section 5.

2. Data

This study focuses on daily-scale, 0.25° (approximately 25 km) precipitation over the conterminous United States (CONUS; see Figure 1). This large geographic extent allows us to robustly demonstrate not only how the CSGD can be used to characterize precipitation

159 and how the CSGD-based error modeling framework can correct for biases and
160 characterize remaining uncertainties, but also how these features vary with climatic and
161 physiographic controls.
162
163 We examine two variants of TMPA (also known as TRMM 3B42) Version 7.0. TMPA
164 merges PMW, active radar, and IR observations from multiple satellites to create a near-
165 global (±50° latitude) rainfall dataset with 3-hourly, 0.25° resolution. The "research"
166 version includes a monthly rain gage-based bias correction and is available approximately
167 two months after realtime. In this study, analyses using this version cover 1998-2014.
168 Several analyses consider TMPA-RT, which is available approximately 8 hours after
169 realtime and only includes a gage-based climatology correction. Such near-realtime
170 analyses cover 2000-2014, since the pre-2000 TRMM orbit precludes near-realtime
171 analysis. "TMPA" is used to refer to the research version and "TMPA-RT" for the near-
172 realtime version. The TRMM satellite ceased operations in April 2015 but the TMPA
173 product is continuing to be produced leveraging other satellites in the constellation.
174 NASA's recent IMERG SMPP was not used in this study, since at the time of writing it
175 was only available for 2014 onward.
176
177 We use the "File A" precipitation forcing from Phase 2 of NASA's National Land Data
178 Assimilation System (NLDAS-2; Xia et al., 2012b, 2012a) as the reference. NLDAS-2
179 precipitation has hourly, 0.125° resolution, disaggregated from daily CPC-Unified gage
180 analysis (Chen et al., 2008; Xie et al., 2007) and features a statistical topographic correction
181 based on the PRISM climatology by Daly et al. (1994). NLDAS-2 was selected rather than

182 the Stage IV bias-corrected radar rainfall dataset that has been used in some SMPP
183 validation studies (AghaKouchak et al., 2011; Qiao et al., 2014) since visual inspection of
184 Stage IV revealed very poor performance in mountainous regions. We have aggregated
185 NLDAS-2 from its hourly 0.125° resolution to the same daily 0.25° resolution as the TMPA
186 data. Thus, the NLDAS-2 precipitation values used in this study are very similar, but not
187 exactly identical, to CPC-Unified, which has been used in several previous SMPP error
188 characterizations (Maggioni et al., 2016, 2014; Tian et al., 2013). The reader is referred to
189 Ferguson and Mocko (2017) for a detailed explanation of the data sources utilized to create
190 the NLDAS-2 precipitation forcing.
191
192 Though there is likely overlap in terms of the rain gages used to create NLDAS-2 and those
193 used to bias-correct the research version of TMPA, the CSGD-based framework does not
194 require strict independence of SMPP and reference data. This study assumes that NLDAS-
195 2 is free of errors, which is of course never the case for any dataset, let alone a continental-
196 scale one such as NLDAS-2. Rain gage undercatch errors in gridded rain gage datasets can
197 be substantial, particularly for snowfall and for extreme rainfall (Adam and Lettenmaier,
198 2003). NLDAS-2 does not use a gage undercatch correction, and thus probably
199 underestimates true precipitation. It should be noted that the monthly gridded rain gage
200 data used to bias correct TMPA does use an undercatch correction. Thorough investigation
201 of the role of gage undercatch errors in satellite precipitation validation is beyond the scope
202 of this study.
203

We also present analyses that utilize surface precipitation rate and vertically integrated total precipitable water (TPW) from Version 2 of the Modern-Era Retrospective analysis for Research and Applications (MERRA-2; Bosilovich et al., 2015; Rienecker et al., 2011) from NASA. MERRA-2 is generated using an atmospheric model that assimilates a range of surface and atmospheric observations including satellite PMW. MERRA-2 outputs have hourly, 0.5° latitude by 0.625° longitude resolution. It is unnecessary to regrid the MERRA-2 datasets to the 0.25° resolution of TMPA for this study, but the same daily temporal resolution is used. Though MERRA-2 provides several surface-level precipitation outputs, including a version primarily based on rain gages, we use model internally-generated precipitation to ensure greater independence from TMPA and NLDAS-2 and to illustrate the value of numerically-generated precipitation and other atmospheric variables for reducing SMPP errors.

The precipitation datasets utilized in this study consider all seasons and precipitation phases (i.e. rain, snow, hail, etc.), represented in terms of depth of liquid water. Determination of precipitation phase is a challenge in gridded precipitation datasets, whether the underlying data come from rain gage networks, satellites, ground-based radar, or numerical models.

We treat data prior to 2014 as the "training period," i.e. used for model parameter estimation as well as error analysis. Data from 2014 is used as "validation," to assess model robustness when used outside of the training period. Though this training period is much

longer than the validation period, this typifies many settings in which an error model might be used, since many reference datasets date at least as far back as most or all SMPPs.

3. Methods

3.1 The CSGD

The two-parameter Gamma distribution has been used in precipitation modeling since at least Das (1955). Like precipitation itself, the Gamma distribution is left-bounded at zero, and can take many possible "shapes," in terms of its density and cumulative distribution function (CDF). Generally, a precipitation process can be modeled in two steps using a total of three parameters. First, the probability of occurrence is modeled via a Bernoulli trial with the "success" parameter equal to the probability of precipitation (POP). Second, the nonzero precipitation magnitude is modeled via the two-parameter Gamma with shape parameter k and scale parameter θ expressed using the distribution mean μ and standard deviation σ by

$$k = \frac{\mu^2}{\sigma^2}, \theta = \frac{\sigma^2}{\mu} \quad (1)$$

The CSGD is an alternative formulation presented in (Scheuerer and Hamill, 2015) in which the CDF is "shifted" left and subsequently left-censored at zero, meaning all negative values are replaced by zero. Thus, the density to the left of zero represents the probability of zero precipitation ($1 - POP$), while the density to the right of zero represents the likelihood of a particular nonzero value. To achieve this, a "shift" parameter δ, $\delta < 0$ is introduced such that, if $F_{k,\theta}$ denotes the CDF of a gamma distribution, then the CDF of the CSGD model is defined by

$$F_{k,\theta,\delta}(x) = \begin{cases} F_{k,\theta}(x-\delta) \text{ for } x \geq 0 \\ 0 \text{ for } x < 0 \end{cases} \quad (2)$$

where x is rainfall depth. In this way, the CSGD eliminates the initial Bernoulli trial from the precipitation modeling process, though the introduction of δ means the total number of parameters remains at three. Thus, while the conventional Gamma distribution has the property that $F_{k,\theta}(0) = 0$ (i.e. the CDF is equal to zero at zero rainfall depth), the CDF of a CSGD has the property $F_{k,\theta,\delta}(0) = 1 - POP$ (see Figure 2). Scheuerer and Hamill (2015) provide details for CSGD parameter estimation based on minimization of the continuous ranked probability score, which essentially minimizes the integrated quadratic distance between the empirical and theoretical CSGD distribution functions.

CDFs for "climatological CSGDs" (to distinguish from conditional CSGDs, described in Section 3.2) are shown for the 0.25° grid cells nearest to Charlotte, North Carolina and Denver, Colorado (top panel of Figure 3). These demonstrate good fit to the empirical CDFs, while highlighting the differences between locations and between TMPA and NLDAS-2.

3.2 CSGD-Based Error Modeling and Correction Framework

The climatological CSGD is insufficient for generating a distribution of estimated "true" precipitation values (or, equivalently, a distribution of SMPP errors) based on a given observation $R_s(t)$ at time t, since the mean $\mu(t)$, standard deviation $\sigma(t)$, and perhaps shift $\delta(t)$ depend on the magnitude of $R_s(t)$. Thus, we use a CSGD-based error modeling framework to reduce systematic SMPP biases, and to model and reduce SMPP random

errors. The framework was first introduced in Scheuerer and Hamill (2015) and further explored in (Báran and Nemoda (2016) for statistical post-processing of ensemble numerical precipitation forecasts. The CSGD-based approach uses a statistical regression model "trained" using a past record of contemporaneous satellite and reference observations. The regression model is then conditioned using a satellite observation for time t to generate "conditional CSGD" parameters $\mu(t)$, $\sigma(t)$, and $\delta(t)$ from the climatological CSGD parameters k, θ, and δ.

In the simplest version, $\mu(t)$ increases linearly with $R_s(t)$ and $\sigma(t)$ increases proportionally to the square root of $\mu(t)$. Allowing $\delta(t)$ to vary offers little benefit and can lead to parameter estimation difficulties (M. Scheuerer, personal comm., February 27, 2017). We will refer to this version as the "linear model," since it models conditional bias linearly with precipitation rate. It has the form

$$\mu(t) = \mu\left(\alpha_2 + \alpha_3 \frac{R_s(t)}{\overline{R_S}}\right) \quad (3)$$

$$\sigma(t) = \alpha_4 \sigma \sqrt{\frac{\mu(t)}{\mu}} \quad (4)$$

$$\delta(t) = \delta \quad (5)$$

where $\overline{\boldsymbol{R_S}}$ denotes the mean of the SMPP time series. Example CDFs of conditional CSGDs are shown in the lower panel of Figure 3 for $R_s(t)$ values of 2.5 and 25 *mm/d* for the 0.25° grid cells nearest to Charlotte, North Carolina and Denver, Colorado. These show that as $R_s(t)$ increases, the probability of the true precipitation being zero decreases (approaching zero for $R_s(t)$=25 *mm/d*) while the probability of higher true values increases. The value

of $\mu(t)$ will always be nonzero and greater than conditional median at time *t*, which will be equal to zero when the conditional POP is less than 0.5.

Scheuerer and Hamill (2015) also present a more complex version that can account for nonlinearity in conditional bias. This model, from now on will be called the nonlinear model, has the form

$$\mu(t) = \frac{\mu}{\alpha_1}\log1p\left[expm1(\alpha_1)\left(\alpha_2 + \alpha_3\frac{R_s(t)}{\overline{R_s}}\right)\right] \quad (6)$$

where $log1p(x) = log\,(1+x)$ and $expm1(x) = exp(x) - 1$.

The regression framework can also accommodate an arbitrary number *n* of additional contemporaneous covariates $C_1(t), C_2(t), \ldots, C_n(t)$ such as TPW, temperature, or humidity from atmospheric observations or simulations. In this case, Equation 3 expands to

$$\mu(t) = \mu\left(\alpha_2 + \alpha_3\frac{R_s(t)}{\overline{R_s}} + \alpha_5\frac{C_1(t)}{\overline{C_1}} + \alpha_6\frac{C_2(t)}{\overline{C_2}} + \cdots + \alpha_{4+n}\frac{C_n(t)}{\overline{C_n}}\right) \quad (7)$$

and $\overline{C_i}$ is the mean of the time series of the *i*th covariate. A similar variant of the nonlinear model (Equation 6) could be written to include covariates. The inclusion of covariates allows for additional information to be introduced to the SMPP-reference intercomparison, allowing the explanation of some of the residual (i.e. random) error. We use the techniques described in Scheuerer and Hamill (2015) to estimate the parameters of the CSGD correction framework.

The models described above are consistent with the notions that satellite errors are multiplicative (Tian et al., 2013) and that error magnitude grows with $R_s(t)$. They bear

312 passing resemblance to the PUSH model of Maggioni et al. (2014), in that the conditional
313 distribution of estimated true precipitation $F_{k(t),\theta(t),\delta(t)}$ given $R_s(t)$ is assumed to be
314 Gamma distributed, though we use the 3-parameter CSGD rather than the conventional 2-
315 parameter Gamma used to model precipitation hits in PUSH. This allows for the possibility
316 of the estimated true precipitation to be zero, even if $R_s(t) > 0$ (i.e. a false alarm) or vice
317 versa (missed precipitation). PUSH, in contrast, accounts for false alarms and misses using
318 separate models, making it impossible to construct a theoretical distribution for estimated
319 true precipitation and involves additional parameters. Like PUSH, the CSGD framework
320 has the advantage of being parametric, which can be helpful in conditions of very low or
321 very high precipitation rates (Gebremichael et al., 2011b; Zhang et al., 2013).

322 # 4. Results and Discussion

323 ## *4.1 CSGD-Based Precipitation Characterization*

324 Estimates of μ, σ, and δ for 1998-2013 for NLDAS-2 and TMPA are compared for every
325 grid cell over CONUS (Figure 4). All three parameters in both TMPA and NLDAS-2
326 exhibit higher values in the eastern United States and the Pacific coastal mountains than in
327 the western United States. This should be expected due to the higher amounts of
328 precipitation in these parts of the country (See Figure 1). TMPA tends to overestimate μ
329 and σ and underestimate δ relative to NLDAS-2 except in the pacific coastal and Rocky
330 Mountains. Differences in μ and σ in the western United States are lower in magnitude,
331 though the relative differences are approximately uniform except for over mountains.
332 Isolated or small clusters of seemingly anomalous parameter values can be seen in TMPA
333 but not in NLDAS-2. Visual inspection shows that these are co-located with water bodies

such as lakes and reservoirs that are known to influence PMW-based precipitation estimates (Tian and Peters-Lidard, 2007).

POP cannot be evaluated directly from Figure 4. Over CONUS, POP for TMPA is more uniform and significantly lower than in NLDAS-2, suggesting that the precipitation detection limits imposed by the satellite sensors or processing algorithms exert strong controls (Figure 5). The TRMM sensor package was designed to detect moderate to heavy rainfall and thus tend to underestimate light precipitation and mixed phase/falling snow. GPM can see a much broader spectrum of precipitation. As with the parameter estimates in Figure 4, anomalous isolated POP values are co-located with water bodies. We do not explore this issue further in this study, but Maggioni et al. (2014) suggest that a minimum detection threshold of 0.25 mm/d may be a reasonable approximation in TMPA and their PUSH error model utilizes this threshold to distinguish between precipitation and non-precipitation. The linear and nonlinear conditional CSGD models described in Section 3.2 do allow for nonzero true precipitation even when $R_s(t) = 0$, and thus the CSGD approach need not explicitly consider detection thresholds.

4.2 Error Modeling using the Conditional CSGD Framework

Before showing CONUS-wide error modeling and correction results using the CSGD framework, we provide a more detailed illustration of the linear and nonlinear models and comparison with the PUSH model from Maggioni et al. (2014) for the 0.25° grid cell nearest to Charlotte, North Carolina (Figure 6). The models and data, including the 1998-2013 training period and 2014 validation period, are shown on both linear (left panels) and

357 logarithmic scales (right panels). For both Charlotte and other locations across CONUS,
358 TMPA tends to overestimate at higher precipitation rates. This overestimation is consistent
359 with previous studies (AghaKouchak et al., 2011; Tian et al., 2009) and may be due to the
360 joint effect of TMPA's monthly bias correction and poor light precipitation detection,
361 which would tend to introduce a high bias in precipitation magnitude (Tian et al., 2009;
362 Wright et al., 2017). However, since NLDAS-2 does not account for gage undercatch, it
363 almost certainly underestimates true heavy precipitation to an unknown degree. Thus, the
364 extent to which TMPA overestimates true precipitation for large events is difficult to assess
365 without a more detailed reference dataset.
366
367 The linear and nonlinear versions of the CSGD-based error model provide good fits to the
368 data for both the training and validation periods, and the nonlinear variant better captures
369 the nonlinearity in conditional bias that is evident in high precipitation. PUSH greatly
370 overestimates conditional bias for high precipitation, and no points fall outside of the lower
371 bound of that model's 95% spread, which is unrealistic given the relatively large sample
372 size. In contrast, approximately 5% of points fall outside of the 95% quantile spread for
373 the CSGD model (note that not all data points are clearly visible in Figure 6, particularly
374 those that fall very close to either axis).
375
376 We evaluate a range of conditional CSGD error model complexities; specifically, models
377 using different versions of Equations 3, 6, and 7 to estimate $\mu(t)$. CONUS-wide evaluation
378 using root-mean-square error (RMSE) from two versions, the linear model without
379 covariates and the nonlinear model with MERRA-2 precipitation, is shown in Figure 7.

Here and in subsequent calculations using CSGD error models, RMSE and other error metrics are computed between NLDAS-2 and the conditional CSGD median. As noted in Section 3.2, the conditional CSGD mean is always nonzero and greater than the median, which for low precipitation rates can be equal to zero. This means that neither the conditional mean nor median are ideal measures of the central tendency, but investigation of a more appropriate summary statistic is beyond the scope of this study. The linear model improves upon the TMPA dataset (i.e. reduces RMSE) except in the Rockies and Pacific coastal mountains, where performance is poor. The nonlinear model with MERRA-2 precipitation offers further improvement, including in these mountainous areas. Reductions in RMSE are greatest in the northern part of the country (particularly the nonlinear model with MERRA-2 precipitation) and in the high-altitude but lower-relief portions of the Intermountain West such as the upper Rio Grande in southern Colorado and northern New Mexico and the Snake River Plain in southern Idaho.

The substantial improvements provided by the nonlinear model with MERRA-2 covariates in the northeastern and northwestern parts of the country are likely attributable to the relatively higher proportion of stratiform precipitation in those regions, which is generally better estimated by atmospheric models than by satellite sensors. The more complex model also improves upon simpler versions in most of the rockies and west coast mountains. Visual inspection of results for a range of models reveal that most of this improvement stems from inclusion of MERRA-2, rather than from the nonlinear model structure (results not shown). Error reductions are associated with the identification and removal of

systematic errors and, in the case of models that include MERRA-2 covariates, some further reduction of random errors.

We compute the RMSE and mean absolute error (MAE) normalized by the mean daily precipitation (henceforth refered to as NRMSE and NMAE, respectively) for each 0.25° grid cell across CONUS for a range of CSGD model configurations. This allows us to compare the relative reduction in errors achieved in various precipitation hydroclimates. Results are then summarized by computing the CONUS-wide median and interquartile range (IQR) of NRMSE and NMAE (Table 1). These nonparametric summary statistics were chosen rather than the mean and standard deviation because in arid parts of the country, normalizing by a daily mean precipitation close to zero can produce spurious results.

The NRMSE and NMAE for the uncorrected TMPA dataset shows slightly increased accuracy for the validation period, relative to the training period, possibly associated with improvements in the number and quality of satellite sensors over the lifetime of TMPA. In contrast, the error statistics for the CSGD models tend be unchanged or slightly worse for the 2014 validation period, though in all cases the validation performance is within 7% of the reference period in terms of RMSE and within 5% in terms of MAE, suggesting relatively robust model performance.

The linear (nonlinear) model improved median NRMSE by 20% (22%) and median NMAE by 17% (19%) for the training period, with similar performance in the validation period.

MERRA-2 covariates improved upon this "baseline" CSGD model performance. The inclusion of MERRA-2 precipitation offers robust improvements to both NRMSE and NMAE (32% and 33%, respectively in the case of the nonlinear version). Inclusion of MERRA-2 TPW alone (i.e. without MERRA-2 precipitation) offers very little improvement in both the linear and nonlinear models. When both MERRA-2 TPW and precipitation are included, neither linear nor nonlinear models show much improvement over when only the precipitation covariate is included. This implies that precipitation from MERRA-2 is a much stronger predictor of true precipitation than TPW. It also suggests that MERRA-2 precipitation and TPW are highly correlated, which is unsurprising.

A linear CSGD error model was tested in which the size of the TMPA and NLDAS-2 samples at each grid cell were expanded by concatenating the data from the eight adjacent grid cells for model fitting. Referred to in Table 1 as "linear with spatial pooling," this model produced similar results to the linear model fitted only to data from individual grid cells ("linear" in Table 1). This has several implications. In complex terrain or near water bodies, precipitation can vary over relatively short distances. In such cases, spatial pooling may create an enlarged sample that does not properly represent precipitation statistics in the grid cell in question. Visual inspection of RMSE maps show similar performance between pooled and unpooled linear CSGD models in the eastern portion of the country, and lower performance using pooling in the mountain west, consistent with this intuition (results not shown). In addition, the value added through spatial pooling is inherently limited if there is substantial spatial correlation in the precipitation estimates and errors between adjacent grid cells. The similar performance between pooled and unpooled models

448 in less varied terrain also implies that the model fitting procedure is relatively robust to
449 small samples.
450
451 We evaluate the relationships between errors in TMPA, as a function of correlation
452 between TMPA and NLDAS-2, before and after applying a nonlinear CSGD model with
453 MERRA-2 precipitation (Figure 8). The influence of land surface elevation, as a proxy for
454 topographic relief, is also evaluated, since this impact is somewhat difficult to assess in
455 Figure 7. Both the absolute values and the variability in NRMSE and NMAE are relatively
456 low for locations with high correlation, while the variability (though not the central
457 tendency) in these statistics increases for locations with lower correlation and there is a
458 relatively weak inverse relationship between error magnitude and correlation between the
459 SMPP and reference. Neither correlation nor elevation appear to be the primary controls
460 on NRMSE or NMAE, even though correlation values for higher-elevation locations tend
461 to be relatively low. It also appears from Figure 8 that similar reductions in NRMSE and
462 NMAE can be achieved regardless of correlation or land surface elevation. Qualitatively
463 similar results were produced with the simpler linear model (not shown).
464
465 Like NRMSE and NMAE, correlation between the uncorrected TMPA and NLDAS-2 is
466 slightly higher in the validation period than the training period, again likely associated with
467 improvements in the quality and number of sensors. Interestingly, linear and nonparametric
468 correlations between corrected SMPP timeseries and NLDAS-2 reduce somewhat when
469 TMPA is fed through a linear CSGD model without covariates, and remain relatively
470 unchanged when a nonlinear model is used instead (Table 2). This may be due to the

limitations of using either the CSGD mean or median and due to the implicit bias adjustment in the CSGD framework. When MERRA-2 precipitation is included as a covariate, however, correlation between the corrected SMPP timeseries and NLDAS-2 increases. This highlights the ability of MERRA-2 covariates (particularly precipitation) to reduce random errors in TMPA.

We also examined the realtime version (TMPA-RT) with several CSGD models (Table 3). NRMSE and NMAE in the original TMPA-RT dataset are 14% larger in terms of NRMSE and 8% larger in terms of NMAE than the research version analyzed previously. Results are qualitatively similar to Table 1, with all CSGD models showing improvement over the uncorrected TMPA-RT dataset, and with the largest improvements coming from the nonlinear model with MERRA-2 precipitation. Likewise, error statistics are generally comparable for the 2014 validation period, showing minimal loss of performance as compared to the training period. The degree of error reduction achieved by the CSGD models is greater using TMPA-RT than TMPA. For example, relative to the uncorrected TMPA-RT, the linear CSGD model reduced NRMSE (NMAE) by 25% (20%), while the same model reduced error for the research version by 20% (17%). Reduction in NRMSE (NMAE) relative to the uncorrected TMPA-RT was as high as 39% (37%) for the nonlinear CSGD with MERRA-2 precipitation. These results are consistent with the notion that error models identify and remove systematic biases, since Maggioni et al. (2016) reported higher systematic errors in TMPA-RT than the research version.

4.3 Parameter Sensitivity

494 The results for the validation period shown in Tables 1 and 2 provide an initial indication
495 that the CSGD framework can be applied outside of the training period. To investigate this
496 issue further, we re-estimate the CSGD parameters for NLDAS-2 and TMPA, as well as
497 the regression parameters for linear version of the conditional CSGD model for each year
498 individually from 1998-2013 and for successively longer time periods (i.e. 1998-1999,
499 1998-2000, etc.) for the grid cell nearest to Charlotte, North Carolina (Figure 9). While
500 parameters vary somewhat from year to year, estimates using longer time periods converge
501 to relatively stable values after several years. Exceptions are the slight downward trend in
502 α_2 and upward trend in α_3. It is well known that the spatial and temporal statistical
503 consistency of precipitation datasets vary according to input data availability, such as the
504 number of rain gages (Hamlet and Lettenmaier, 2005) or the quality and type of satellite
505 sensor (Cho and Chun, 2008). The trends in α_2 and α_3 are consistent with improvement in
506 precipitation estimation in TMPA (i.e. reduction in the weight given to the regression
507 intercept and increase in weight given to R_s). Parameters for the nonlinear model and for
508 other locations are similarly stable over time (results not shown).
509
510 These results suggest that the continuous ranked probability score-based parameter
511 estimation procedure for the climatological CSGD and the conditional CSGD regression
512 framework is relatively efficient with respect to data requirements, and that several years
513 of coincident reference data may be sufficient. It would be worthwhile to evaluate this issue
514 using error metrics such as RMSE or MAE. We leave this as a topic of future work, though
515 it is worth noting that (Scheuerer and Hamill, 2015) found relatively poor conditional

516 CSGD performance with a one year training sample but good performance with modest
517 increases in training record length.

518 ## 5. Summary and Discussion

519 Using the censored shifted gamma distribution (CSGD), we characterize the climatology
520 of daily precipitation over CONUS of TMPA, a satellite multisensor precipitation product
521 (SMPP) and NLDAS-2, a reference (i.e. rain gage-based) dataset. We also use a conditional
522 CSGD error modeling framework to quantify and reduce errors in TMPA. The CSGD
523 describes both precipitation occurrence and magnitude, and reveals significant differences
524 between TMPA and NLDAS-2 including poor satellite-based estimation over inland water
525 bodies and mountainous regions. The CSGD-based error modeling framework considers
526 errors both in the detection and magnitude of precipitation and can model systematic bias
527 either as a linear or nonlinear function of precipitation rate. Both versions perform better
528 than an existing error model from Maggioni et al. (2014) over a wide range of precipitation
529 magnitudes for daily precipitation.
530
531 The framework suffers most in areas of high topographic relief (though not necessarily in
532 areas of high elevation). Error reduction at a specific location depends on the relative
533 balance of systematic and random error in the SMPP at that location. Preliminary analyses
534 demonstrate that parameter estimation of both the CSGD and the CSGD-based error
535 framework are relatively insensitive to record length for periods of record longer than
536 several years.
537

In addition, we show that errors in TMPA can be reduced by incorporating covariates from MERRA-2 atmospheric reanalysis, despite its relatively coarse resolution. This is the first study that we are aware of in which the potential benefits of merging numerical weather prediction and SMPP is explored quantitatively. Precipitation from MERRA-2 offers robust increases in performance, particularly in mountainous areas, while MERRA-2 precipitable water provides little improvement. The improvements offered by MERRA-2 appear to be due to the better performance of numerical models relative to satellite-based instruments, in resolving stratiform precipitation. Other numerical weather models that have higher resolution or that assimilate more independent observations would likely provide additional improvement.

It should be emphasized that precipitation error models can only isolate and thus remove systematic errors. The errors remaining after the removal of systematic bias, i.e. the random errors, can be described statistically but not reduced or eliminated. The variability in these residuals can only be explained via the inclusion of additional information. Except for models that include MERRA-2 covariates, therefore, the error reductions shown throughout Section 4.2 stem solely from the identification and removal of systematic errors. MERRA-2 covariates can explain some amount of residual (i.e. random) error, as evidenced by the further reductions in errors and increased correlations.

The error reduction achieved in this study is generally consistent with the levels of systematic error found over the eastern United States at the same spatial and temporal resolution by Maggioni et al. (2016), though more work is needed to reconcile

561 discrepancies between the degree of systematic bias shown here and shown by those
562 authors in the mountain west. Also consistent with (Maggioni et al., 2016), improvements
563 in TMPA-RT were relatively greater than for the gage-corrected non-realtime version,
564 suggesting that the CSGD approach has particular advantages for near-realtime
565 applications. The CSGD approach, coupled with realtime numerical weather prediction
566 estimates such as those generated using NASA's GEOS-5 (Rienecker et al., 2008), offer a
567 pathway to improve the accuracy of near-realtime SMPP, and for parameterizing remaining
568 random errors.
569
570 Certain relevant issues were not explored in this study. Maggioni et al. (2014) concluded
571 that seasonally varying model parameters offered no major advantage in their error model,
572 and our initial investigations into seasonality, which are omitted here in the interest of
573 brevity, confirm this. Errors in the NLDAS-2 reference data, including due to rain gage
574 undercatch, were not considered and can be significant, particularly in the cold season and
575 in steep terrain.
576
577 Many applications, such as hydrologic modeling, can require subdaily precipitation inputs.
578 SMPP errors in magnitude grow with increasing resolution. The autocorrelation of daily
579 precipitation is relatively low, but increases as temporal resolution becomes finer. Thus,
580 generating a realistic high-resolution timeseries of precipitation using the CSGD approach
581 or other error models requires consideration of this autocorrelation. The same is true for
582 generating spatially-correlated precipitation fields.
583

584 One key challenge with the CSGD framework, and precipitation error modeling more
585 generally, is transferability to regions that lack reference data. This issue requires
586 significant further effort, but several previous studies have shown promise (Gebregiorgis
587 and Hossain, 2014, 2013). The CSGD framework would be strong candidate for such
588 efforts, due to the relatively simple structure, robust performance, and the ability to include
589 relevant atmospheric variables from numerical weather prediction, which may potentially
590 be even more useful in data-limited settings. Resolving such issues would constitute a
591 major step toward quantifying and reducing errors in satellite precipitation estimates and
592 helping users to better understand the implications of remaining irreducible random errors.

593 ## Acknowledgements

594 This work was supported by NASA's Precipitation Measurement Mission Grant Number
595 NNX16AH72G and the Wisconsin Alumni Research Foundation and used computing
596 resources and assistance from the UW-Madison Center For High Throughput Computing
597 (CHTC). CHTC is supported by UW-Madison, the Advanced Computing Initiative, the
598 Wisconsin Alumni Research Foundation, the Wisconsin Institutes for Discovery, and the
599 National Science Foundation, and is an active member of the Open Science Grid, which is
600 supported by the National Science Foundation and the U.S. Department of Energy. We
601 thank Dr. Michael Scheuerer at the University of Colorado Cooperative Institute for
602 Research in Environmental Sciences and NOAA Earth System Research Laboratory for
603 providing advice and his original CSGD code. We also thank Dr. Viviana Maggioni at
604 George Mason University for direction regarding PUSH.

605 **References**

606 Adam, J.C., Lettenmaier, D.P., 2003. Adjustment of global gridded precipitation for
607 systematic bias. J. Geophys. Res. 108, 4257. doi:10.1029/2002JD002499
608 AghaKouchak, A., Bárdossy, A., Habib, E., 2010. Conditional simulation of remotely
609 sensed rainfall data using a non-Gaussian v-transformed copula. Adv. Water
610 Resour. 33, 624–634. doi:10.1016/j.advwatres.2010.02.010
611 AghaKouchak, A., Behrangi, A., Sorooshian, S., Hsu, K., Amitai, E., 2011. Evaluation of
612 satellite-retrieved extreme precipitation rates across the central United States. J.
613 Geophys. Res. Atmospheres 116, n/a--n/a. doi:10.1029/2010JD014741
614 Anders, A.M., Roe, G.H., Durran, D.R., Minder, J.R., 2007. Small-Scale Spatial
615 Gradients in Climatological Precipitation on the Olympic Peninsula. J.
616 Hydrometeorol. 8, 1068–1081. doi:10.1175/JHM610.1
617 Báran, S., Nemoda, D., 2016. Censored and shifted gamma distribution based EMOS
618 model for probabilistic quantitative precipitation forecasting. Environmetrics 27,
619 280–292. doi:10.1002/env.2391
620 Behrangi, A., Khakbaz, B., Jaw, T.C., AghaKouchak, A., Hsu, K., Sorooshian, S., 2011.
621 Hydrologic evaluation of satellite precipitation products over a mid-size basin. J.
622 Hydrol. 397, 225–237. doi:10.1016/j.jhydrol.2010.11.043
623 Bosilovich, M., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R.,
624 Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L.,
625 Vikhliaev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson,
626 S., Reale, O., Schubert, S.D., Suarez, M., 2015. MERRA-2 : Initial Evaluation of
627 the Climate. NASA Tech. Rep. Ser. Glob. Model. Data Assim. 43.
628 Chen, M., Shi, W., Xie, P., Silva, V.B.S., Kousky, V.E., Wayne Higgins, R., Janowiak,
629 J.E., 2008. Assessing objective techniques for gauge-based analyses of global
630 daily precipitation. J. Geophys. Res. Atmospheres 113, n/a-n/a.
631 doi:10.1029/2007JD009132
632 Cho, H.K., Chun, H.Y., 2008. Impacts on the TRMM data due to orbit boost in the
633 spectral domain. Geophys. Res. Lett. 35. doi:10.1029/2007GL032320
634 Ciach, G.J., Krajewski, W.F., Villarini, G., 2007. Product-Error-Driven Uncertainty
635 Model for Probabilistic Quantitative Precipitation Estimation with NEXRAD
636 Data. J. Hydrometeorol. 8, 1325–1347. doi:10.1175/2007JHM814.1
637 Ciach, G.J., Morrissey, M.L., Krajewski, W.F., 2000. Conditional Bias in Radar Rainfall
638 Estimation. J Appl Meteor 39, 1941–1946. doi:10.1175/1520-
639 0450(2000)039<1941:CBIRRE>2.0.CO;2
640 Daly, C., Neilson, R.P., Phillips, D.L., 1994. A Statistical-Topographic Model for
641 Mapping Climatological Precipitation over Mountainous Terrain. J. Appl.
642 Meteorol. 33, 140–158. doi:10.1175/1520-
643 0450(1994)033<0140:ASTMFM>2.0.CO;2
644 Das, S.C., 1955. The fitting of truncated type III curves to daily rainfall data. Aust. J.
645 Phys. 8, 298–304.
646 Ebert, E.E., Janowiak, J.E., Kidd, C., 2007. Comparison of near-real-time precipitation
647 estimates from satellite observations and numerical models. Bull. Am. Meteorol.
648 Soc. 88, 47–64. doi:10.1175/BAMS-88-1-47

649 Ferguson, C.R., Mocko, D.M., 2017. Diagnosing an artificial trend in NLDAS-2
650 afternoon precipitation. J. Hydrometeorol. JHM-D-16-0251.1. doi:10.1175/JHM-
651 D-16-0251.1
652 Ferraro, R.R., Peters-Lidard, C.D., Hernandez, C., Turk, F.J., Aires, F., Prigent, C., Lin,
653 X., Boukabara, S.-A., Furuzawa, F.A., Gopalan, K., Harrison, K.W., Karbou, F.,
654 Li, L., Liu, C., Masunaga, H., Moy, L., Ringerud, S., Skofronick-Jackson, G.M.,
655 Tian, Y., Wang, N.-Y., 2013. An Evaluation of Microwave Land Surface
656 Emissivities Over the Continental United States to Benefit GPM-Era Precipitation
657 Algorithms. IEEE Trans. Geosci. Remote Sens. 51, 378–398.
658 doi:10.1109/TGRS.2012.2199121
659 Gebregiorgis, A., Hossain, F., 2014. Making Satellite Precipitation Data Work for the
660 Developing World. Geosci. Remote Sens. Mag. IEEE.
661 doi:10.1109/MGRS.2014.2317561
662 Gebregiorgis, A.S., Hossain, F., 2013. Understanding the Dependence of Satellite
663 Rainfall Uncertainty on Topography and Climate for Hydrologic Model
664 Simulation. IEEE Trans. Geosci. Remote Sens. 51, 704–718.
665 doi:10.1109/TGRS.2012.2196282
666 Gebremichael, M., Liao, G., Yan, J., 2011a. Nonparametric error model for a high
667 resolution satellite rainfall product. Water Resour. Res. 47, n/a-n/a.
668 doi:10.1029/2010WR009667
669 Gebremichael, M., Liao, G.-Y., Yan, J., 2011b. Nonparametric error model for a high
670 resolution satellite rainfall product. Water Resour. Res. 47, n/a-n/a.
671 doi:10.1029/2010WR009667
672 Germann, U., Berenguer, M., Sempere-Torres, D., Zappa, M., 2009. REAL-Ensemble
673 radar precipitation estimation for hydrology in a mountainous region. Q. J. R.
674 Meteorol. Soc. 135, 445–456. doi:10.1002/qj.375
675 Hamlet, A.F., Lettenmaier, D.P., 2005. Production of Temporally Consistent Gridded
676 Precipitation and Temperature Fields for the Continental United States. J.
677 Hydrometeorol. 6, 330–336. doi:10.1175/JHM420.1
678 Hossain, F., Anagnostou, E.N., 2006. A two-dimensional satellite rainfall error model.
679 IEEE Trans. Geosci. Remote Sens. 44, 1511–1522.
680 doi:10.1109/TGRS.2005.863866
681 Huffman, G.J., Adler, R.F., Bolvin, D.T., Nelkin, E.J., 2010. The TRMM Multi-satellite
682 Precipitation Analysis (TMPA), in: Hossain, F., Gebremichael, M. (Eds.),
683 Satellite Rainfall Applications for Surface Hydrology. Springer Verlag, pp. 3–22.
684 Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R.J., Xie, P., 2014.
685 Algorithm Theoretical Basis Document (ATBD) Version 4- NASA Global
686 Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM
687 (IMERG), PMM Website.
688 Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y.,
689 Bowman, K.P., Stocker, E.F., 2007. The TRMM Multisatellite Precipitation
690 Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation
691 Estimates at Fine Scales. J. Hydrometeorol. 8, 38–55. doi:10.1175/JHM560.1
692 Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: A method that
693 produces global precipitation estimates from passive microwave and infrared data
694 at high spatial and temporal resolution. J. Hydrometeorol. 5, 487–503.

695 Maggioni, V., Sapiano, M.R.P., Adler, R.F., 2016. Estimating Uncertainties in High-
696 Resolution Satellite Precipitation Products: Systematic or Random Error? J.
697 Hydrometeorol. 17, 1119–1129. doi:10.1175/JHM-D-15-0094.1
698 Maggioni, V., Sapiano, M.R.P., Adler, R.F., Tian, Y., Huffman, G.J., 2014. An Error
699 Model for Uncertainty Quantification in High-Time-Resolution Precipitation
700 Products. J Hydrometeor 15, 1274–1292. doi:10.1175/JHM-D-13-0112.1
701 Newman, A.J., Clark, M.P., Craig, J., Nijssen, B., Wood, A., Gutmann, E., Mizukami, N.,
702 Brekke, L., Arnold, J.R., 2015. Gridded Ensemble Precipitation and Temperature
703 Estimates for the Contiguous United States. J. Hydrometeorol. 16, 2481–2500.
704 doi:10.1175/JHM-D-15-0026.1
705 Qiao, L., Hong, Y., Chen, S., Zou, C.B., Gourley, J.J., Yong, B., 2014. Performance
706 assessment of the successive Version 6 and Version 7 TMPA products over the
707 climate-transitional zone in the southern Great Plains, USA. J. Hydrol. 513, 446–
708 456. doi:10.1016/j.jhydrol.2014.03.040
709 Rienecker, M., Suarez, M., Todling, R., 2008. The GEOS-5 Data Assimilation System-
710 Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0. NASATM–2008–104606
711 Vol 27 Tech. 27.
712 Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E.,
713 Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.K., Bloom, S., Chen, J.,
714 Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi,
715 R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R.,
716 Robertson, F.R., Ruddick, A.G., Sienkiewicz, M., Woollen, J., 2011. MERRA:
717 NASA's modern-era retrospective analysis for research and applications. J. Clim.
718 24, 3624–3648. doi:10.1175/JCLI-D-11-00015.1
719 Ringerud, S., Kummerow, C., Peters-Lidard, C., Tian, Y., Harrison, K., 2014. A
720 Comparison of Microwave Window Channel Retrieved and Forward-Modeled
721 Emissivities Over the U.S. Southern Great Plains. IEEE Trans. Geosci. Remote
722 Sens. 52, 2395–2412. doi:10.1109/TGRS.2013.2260759
723 Sarachi, S., Hsu, K., Sorooshian, S., 2015. A Statistical Model for the Uncertainty
724 Analysis of Satellite Precipitation Products. J. Hydrometeorol. 16, 2101–2117.
725 doi:10.1175/JHM-D-15-0028.1
726 Scheuerer, M., Hamill, T.M., 2015. Statistical Post-Processing of Ensemble Precipitation
727 Forecasts by Fitting Censored, Shifted Gamma Distributions. Mon. Weather Rev.
728 150901110234004. doi:10.1175/MWR-D-15-0061.1
729 Shige, S., Kida, S., Ashiwake, H., Kubota, T., Aonashi, K., 2013. Improvement of TMI
730 rain retrievals in mountainous areas. J. Appl. Meteorol. Climatol. 52, 242–254.
731 doi:10.1175/JAMC-D-12-074.1
732 Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H. V., Imam, B., Braithwaite, D., 2000.
733 Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall.
734 Bull. Am. Meteorol. Soc. 81, 2035–2046. doi:10.1175/1520-
735 0477(2000)081<2035:EOPSSE>2.3.CO;2
736 Tian, Y., Huffman, G.J., Adler, R.F., Tang, L., Sapiano, M., Maggioni, V., Wu, H., 2013.
737 Modeling errors in daily precipitation measurements: Additive or multiplicative?
738 Geophys. Res. Lett. 40, 2060–2065. doi:10.1002/grl.50320

739 Tian, Y., Peters-Lidard, C.D., 2007. Systematic anomalies over inland water bodies in
740 satellite-based precipitation estimates. Geophys. Res. Lett. 34, L14403.
741 doi:10.1029/2007GL030787
742 Tian, Y., Peters-Lidard, C.D., Eylander, J.B., Joyce, R.J., Huffman, G.J., Adler, R.F.,
743 Hsu, K., Turk, F.J., Garcia, M., Zeng, J., 2009. Component analysis of errors in
744 satellite-based precipitation estimates. J. Geophys. Res. 114, D24101.
745 doi:10.1029/2009JD011949
746 Wright, D.B., Mantilla, R., Peters-Lidard, C.D., 2017. A remote sensing-based tool for
747 assessing rainfall-driven hazards. Environ. Model. Softw. 90, 34–54.
748 doi:10.1016/j.envsoft.2016.12.006
749 Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., Alonge, C., Wei, H.,
750 Meng, J., Livneh, B., Duan, Q., Lohmann, D., 2012a. Continental-scale water and
751 energy flux analysis and validation for North American Land Data Assimilation
752 System project phase 2 (NLDAS-2): 2. Validation of model-simulated
753 streamflow. J. Geophys. Res. 117, D03110. doi:10.1029/2011JD016051
754 Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C.,
755 Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan,
756 Y., Mocko, D., 2012b. Continental-scale water and energy flux analysis and
757 validation for the North American Land Data Assimilation System project phase 2
758 (NLDAS-2): 1. Intercomparison and application of model products. J. Geophys.
759 Res. 117, D03109. doi:10.1029/2011JD016048
760 Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., Yang, S., 2007. A
761 Gauge-Based Analysis of Daily Precipitation over East Asia. J. Hydrometeorol. 8,
762 607. doi:10.1175/JHM583.1
763 Yan, J., Gebremichael, M., 2009. Estimating actual rainfall from satellite rainfall
764 products. Atmospheric Res. 92, 481–488. doi:10.1016/j.atmosres.2009.02.004
765 Zhang, Y., Habib, E., Kuligowski, R.J., Kim, D., 2013. Joint distribution of multiplicative
766 errors in radar and satellite QPEs and its use in estimating the conditional
767 exceedance probability. Adv. Water Resour. 59, 133–145.
768 doi:10.1016/j.advwatres.2013.06.004
769

770

771

772

773

774

Tables

Table 1: Median of CONUS-wide NRMSE and NMAE for TMPA vs. NLDAS-2 and for a range of CSGD error models. Values in parentheses give the interquartile range (IQR; i.e, 25^{th}-75^{th} percentiles). The models are fit to the 1998-2013 time period, while 2014 is reserved for validation.

	NRMSE [-]		**NMAE [-]**	
CSGD Error Model	**1998-2013**	**2014**	**1998-2013**	**2014**
Uncorrected TMPA	2.73 (2.27, 3.25)	2.54 (2.12, 3.14)	0.98 (0.87, 1.11)	0.92 (0.81, 1.05)
Linear	2.19 (1.86, 2.74)	2.25 (1.89, 2.84)	0.81 (0.74, 0.89)	0.80 (0.73, 0.89)
Linear with spatial pooling	2.20 (1.87, 2.73)	2.26 (1.89, 2.85)	0.81 (0.74, 0.89)	0.80 (0.73, 0.89)
Nonlinear	2.14 (1.82, 2.70)	2.22 (1.84, 2.83)	0.79 (0.72, 0.88)	0.79 (0.72, 0.88)
Linear with MERRA-2 precipitation	1.88 (1.55, 2.36)	1.99 (1.60, 2.58)	0.67 (0.60, 0.75)	0.69 (0.61, 0.78)
Linear with MERRA-2 TPW	2.17 (1.83, 2.71)	2.22 (1.85, 2.82)	0.79 (0.72, 0.88)	0.79 (0.71, 0.88)
Linear with MERRA-2 precipitation and TPW	1.87 (1.54, 2.36)	1.98 (1.59, 2.56)	0.67 (0.60, 0.75)	0.68 (0.61, 0.78)
Nonlinear with MERRA-2 precipitation	1.85 (1.53, 2.33)	1.97 (1.58, 2.55)	0.66 (0.59, 0.74)	0.69 (0.61, 0.77)
Nonlinear with MERRA-2 TPW	2.13 (1.80, 2.69)	2.21 (1.82, 2.80)	0.78 (0.71, 0.87)	0.78 (0.71, 0.87)
Nonlinear with MERRA-2 precipitation and TPW	1.84 (1.52, 2.33)	1.97 (1.57, 2.55)	0.66 (0.58, 0.74)	0.69 (0.61, 0.78)

803 Table 2: CONUS-wide median and IQR for Pearson and Spearman correlation coefficients
804 for TMPA vs. NLDAS-2 and for a range of CSGD error models. The models are fit to the
805 1998-2013 period, while 2014 is reserved for validation.

	Pearson Correlation		**Spearman Correlation**	
CSGD Error Model	**1998-2013**	**2014**	**1998-2013**	**2014**
Uncorrected TMPA	0.65 (0.53, 0.71)	0.67 (0.56, 0.74)	0.53 (0.44, 0.62)	0.58 (0.49, 0.65)
Linear	0.63 (0.51, 0.69)	0.65 (0.53, 0.73)	0.52 (0.43, 0.60)	0.56 (0.46, 0.63)
Nonlinear	0.65 (0.53, 0.71)	0.67 (0.56, 0.74)	0.53 (0.44, 0.61)	0.57 (0.47, 0.63)
Linear with MERRA-2 precipitation	0.74 (0.68, 0.79)	0.75 (0.67, 0.81)	0.70 (0.62, 0.75)	0.72 (0.65, 0.78)
Nonlinear with MERRA-2 precipitation	0.75 (0.70, 0.80)	0.76 (0.68, 0.81)	0.71 (0.64, 0.76)	0.73 (0.66, 0.78)

806

807

808 Table 3: As per Table 1, but using TMPA-RT and with a reduced set of CSGD error
809 models.

	NRMSE [-]		**NMAE [-]**	
CSGD Error Model	**1998-2013**	**2014**	**1998-2013**	**2014**
Uncorrected TMPA-RT	3.10 (2.29, 4.24)	3.08 (2.37, 4.09)	1.06 (0.87, 1.38)	1.05 (0.90, 1.29)
Linear	2.32 (1.93, 3.00)	2.38 (1.96, 3.06)	0.84 (0.76, 0.94)	0.84 (076, 0.93)
Nonlinear	2.25 (1.87, 2.94)	2.32 (1.89, 3.01)	0.82 (0.73, 0.92)	0.83 (0.74, 0.92)
Linear with MERRA-2 precipitation	1.91 (1.56, 2.49)	2.05 (1.63, 2.68)	0.68 (0.60, 0.78)	0.71 (0.62, 0.81)
Nonlinear with MERRA-2 precipitation	1.88 (1.55, 2.44)	2.01 (1.60, 2.64)	0.67 (0.59, 0.76)	0.70 (0.62, 0.80)

810

Figures

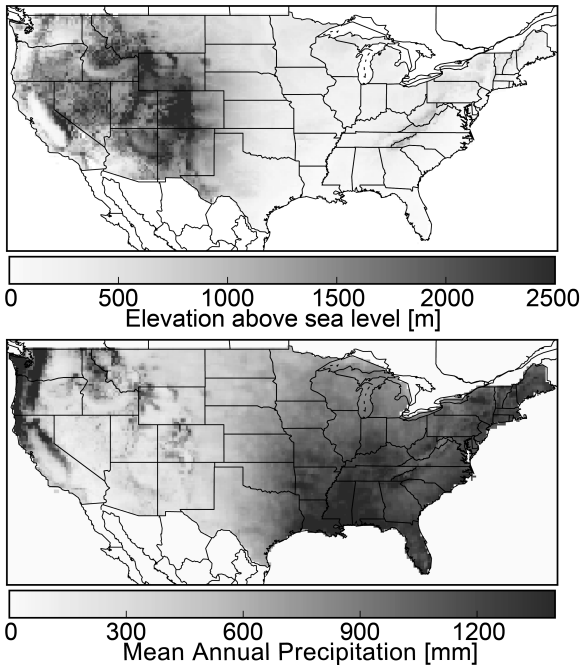


Figure 1: CONUS study area land surface elevation (top) and mean annual precipitation from NLDAS-2 (bottom).

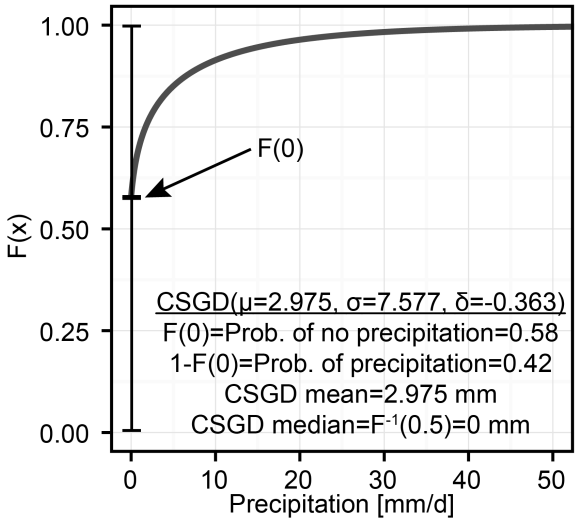
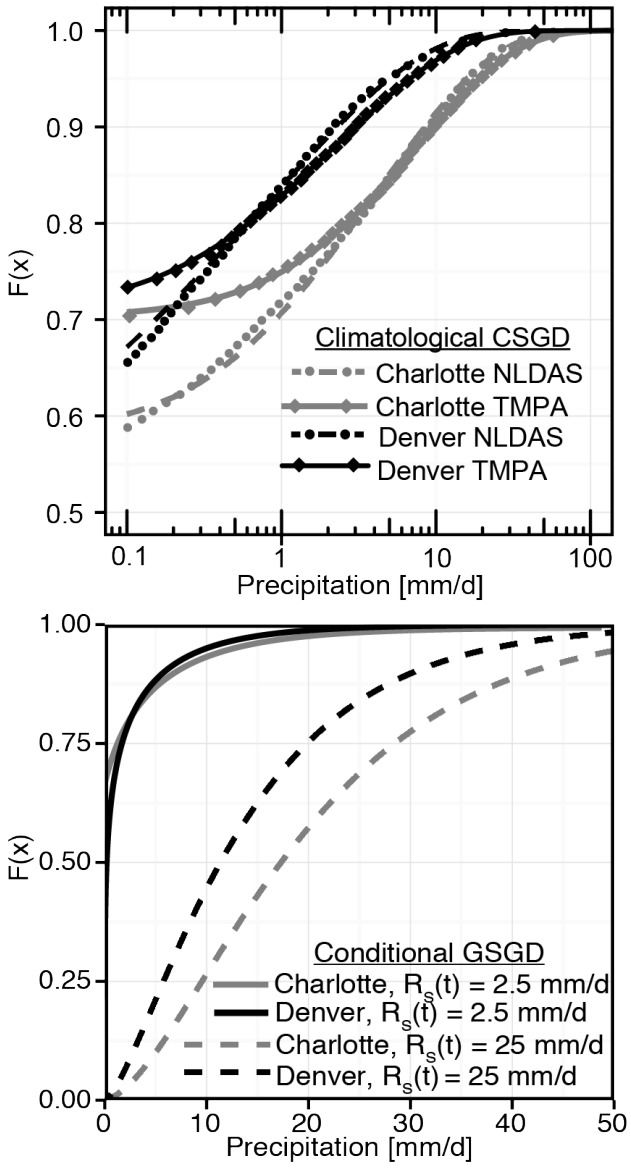


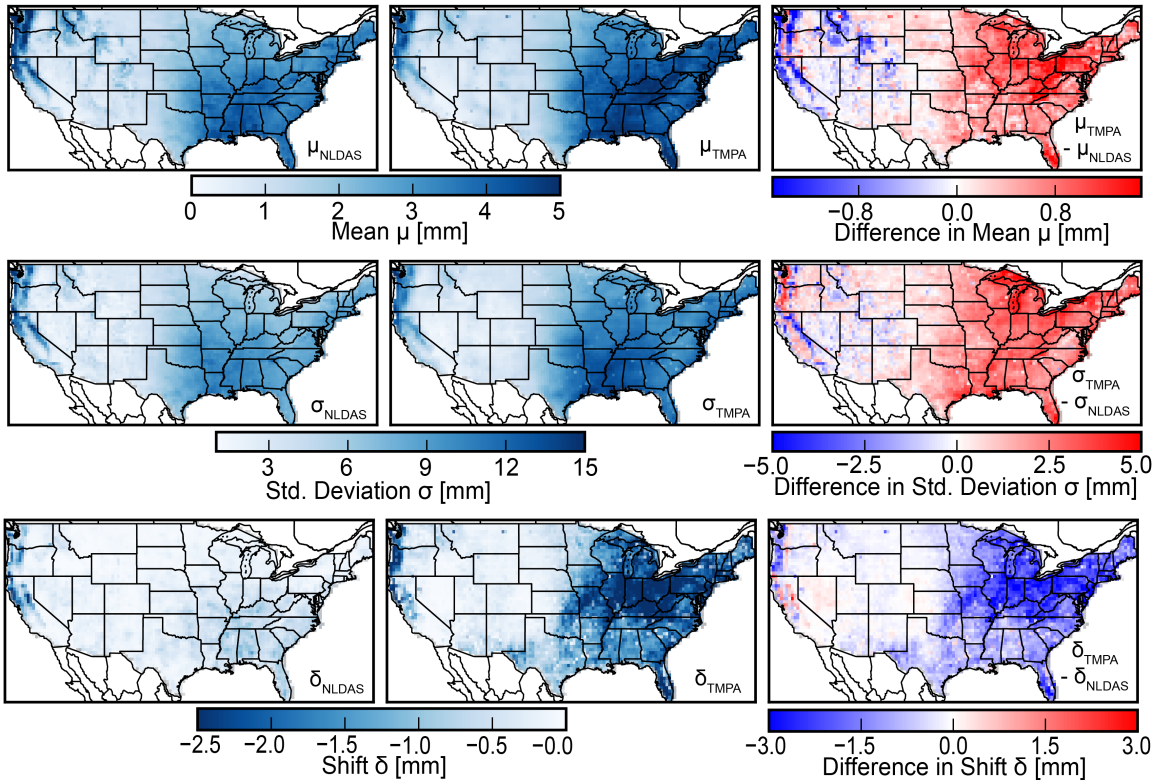
Figure 2: CDF for an arbitrary CSGD distribution. Note that the CDF fully describes both the probability of zero and non-zero precipitation, as well as precipitation intensity.



854

855 Figure 3: Top panel—empirical CDFs (markers) and CSGD theoretical CDFs (lines) for
856 NLDAS-2 and TMPA for Charlotte, North Carolina and Denver, Colorado. A log scale is
857 used for rainfall to improve readability. Bottom panel—conditional CSGD theoretical
858 CDFs generated using the linear model described in Section 3 for $R_s(t)$= 2.5 and 25 mm/d.

859

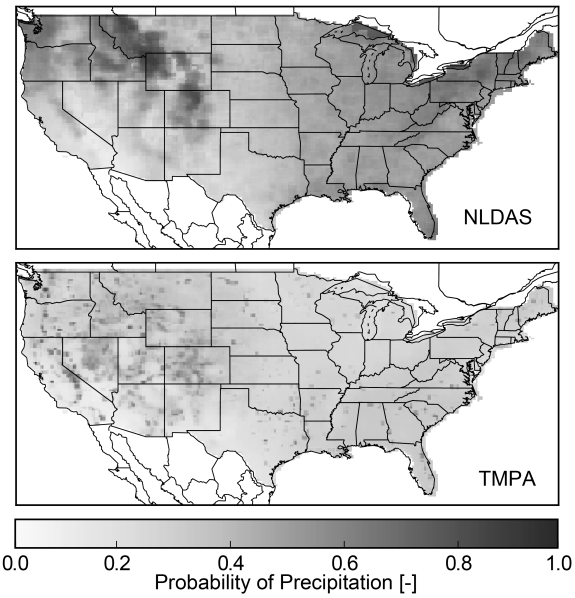


860

861 Figure 4: Climatological CSGD parameters μ, σ, and δ for the 1998-2013 period for

862 NLDAS-2 (left), TMPA (middle), and the difference (right).

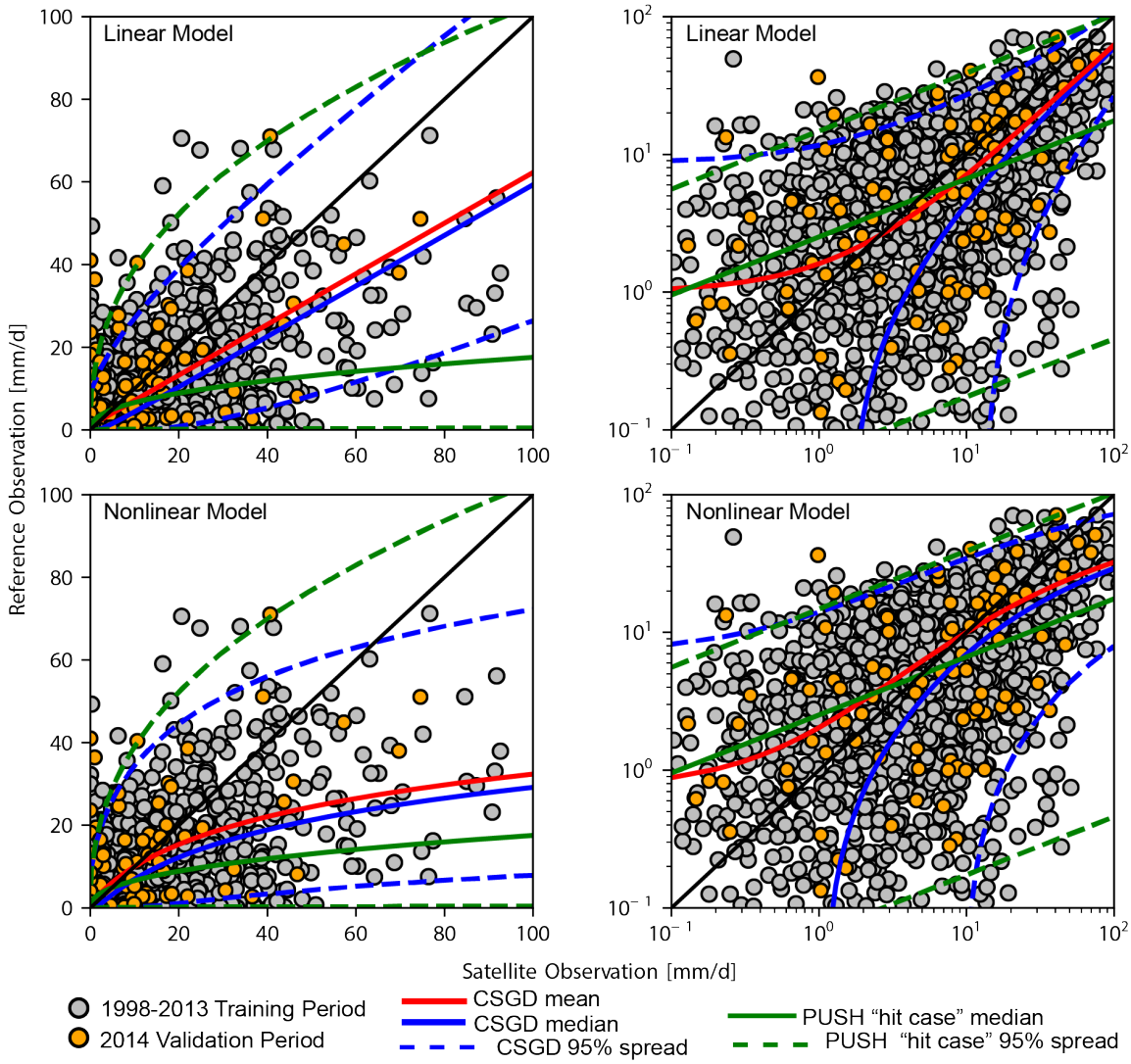
863
864
865
866
867



868

869 Figure 5: Probability of precipitation for the 1998-2013 period using NLDAS-2 (top) and

870 TMPA (bottom).

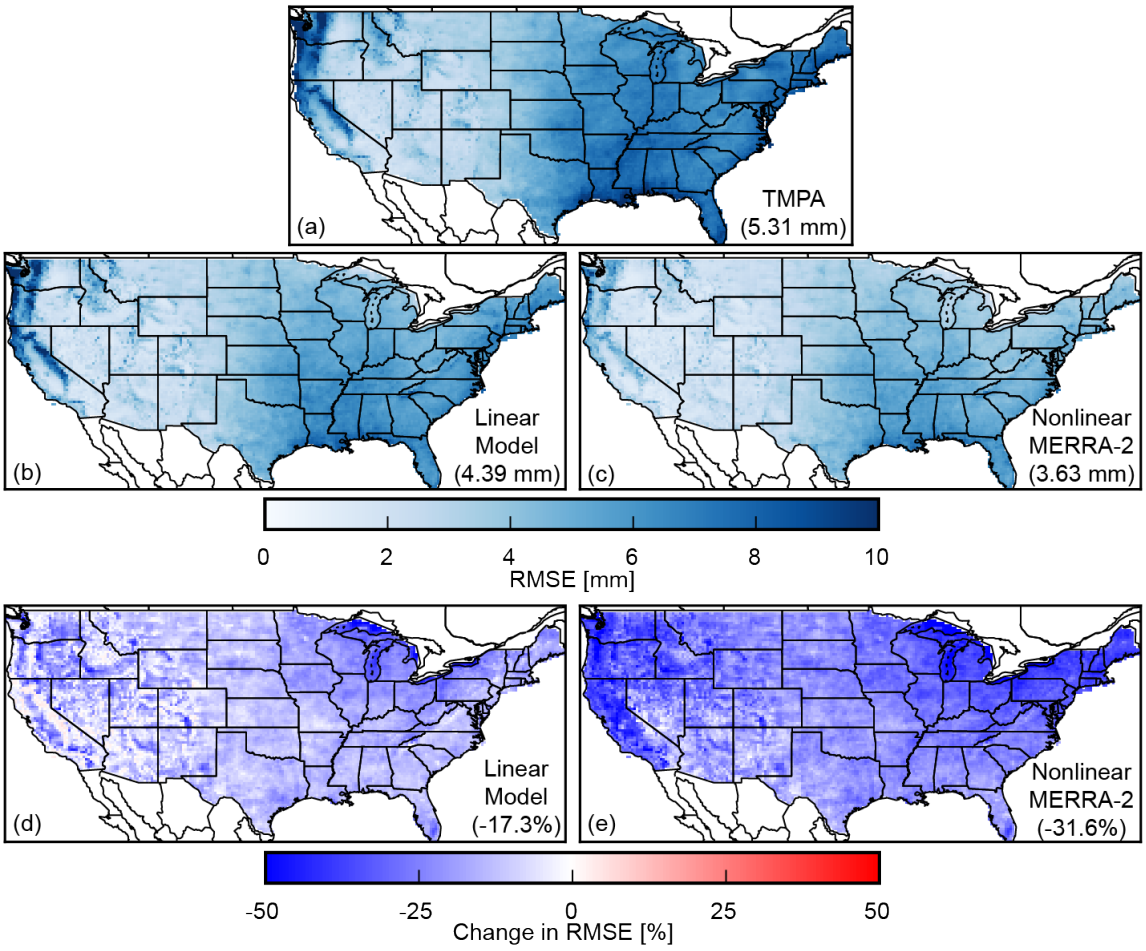


871

872 Figure 6: Linear (top panels) and nonlinear (bottom panels) conditional CSGD models for
873 the 0.25° grid cell nearest to Charlotte, North Carolina compared with observations and
874 PUSH model for 1998-2013 training period (grey dots) and 2014 validation period (orange
875 dots). The sample data and models are shown in the left and right panels but the axes are
876 linear (left panels) and logarithmic (right panels).

877

878



879

880 Figure 7: Top and middle panels—all-season RMSE for 1998-2013, computed relative to
881 NLDAS-2 reference: (a) research version of TMPA; (b) linear model; (c) nonlinear model
882 with MERRA-2 precipitation. Bottom panels—percentage change in RMSE relative to
883 TMPA results in panel (a): (d) linear model; (e) nonlinear model with MERRA-2
884 precipitation. Inset values in parentheses are the means of all grid cells in CONUS.

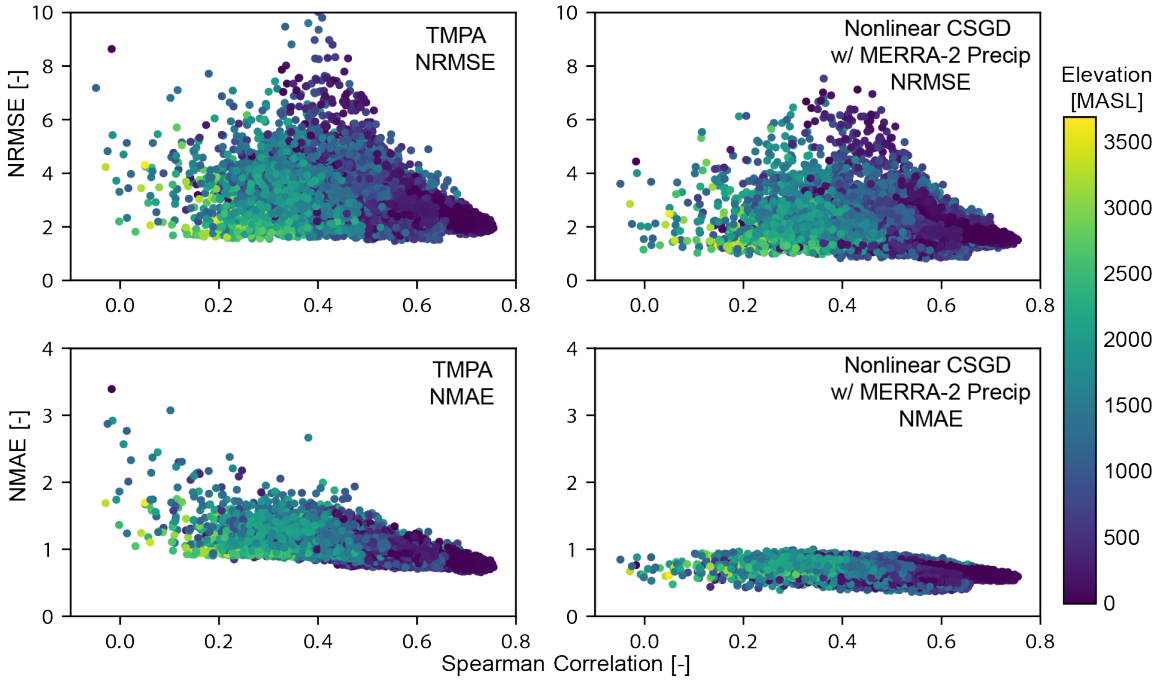
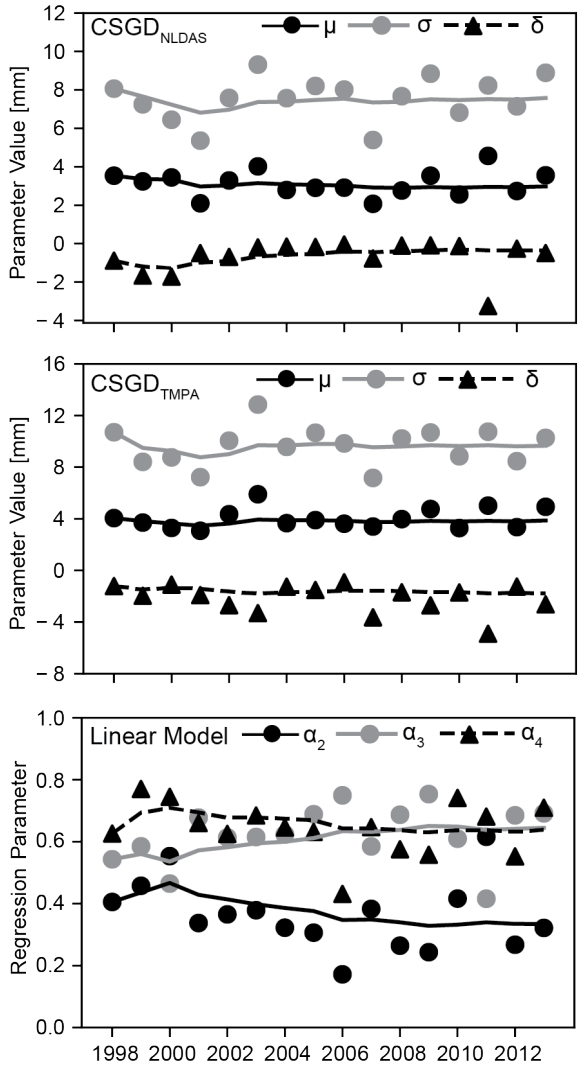


Figure 8: NRMSE (top panels) and NMAE (bottom panels) as a function of Spearman correlation coefficient for every 0.25° in the CONUS study domain. Left panels show results for the TMPA dataset for 1998-2013; right panels show results for the nonlinear CSGD model with MERRA-2 precipitation. Point colors indicate average land surface elevation in the grid cell.



897

898 Figure 9: Parameter estimates as a function of precipitation record length from 1998-2013
899 for the 0.25° grid cell nearest to Charlotte, North Carolina. Top: CSGD for NLDAS-2;
900 middle: CSGD for TMPA. Bottom: regression parameters for linear model. Markers
901 indicate parameter estimates based on that individual year of data, while the lines indicate
902 parameter estimates based on data from 1998 to that year.